

文章编号 1004-924X(2025)13-2108-16

基于光照引导的三阶段低光照图像增强

张易航, 钟 寒*

(中国人民公安大学 信息安全学院, 北京 100038)

摘要:针对在低光照条件下拍摄图像噪声严重、色彩偏差、伪影等问题,提出了一个基于光照引导的三阶段低光照图像增强算法(IG-TSNet)。该算法将傅里叶频域在捕捉图像全局信息的优势与Transformer在空域的远程依赖建模优势结合起来,并采用光照信息指导增强全过程。IG-TSNet由三个阶段级联组成,在像素增强阶段,通过自适应参数调整机制增强图像的全局表征;在傅里叶重建阶段,使用光照先验对幅度和相位谱进行双通道优化,实现图像全局重建;在交叉注意力融合阶段,设计了一个内嵌轻量级交叉注意力机制的双路径U型结构网络,对经过傅里叶重建的输出与光照分支获得的光照引导信息进行交叉注意力融合。最后,在6个LLIE基准数据集上进行了算法评估。实验结果表明,IG-TSNet可以有效提高欠曝光区域的亮度,可靠抑制噪声不会引入斑块和伪影,并稳健保持图像色准。IG-TSNet在9项指标上达到最优。在3个配对数据集上的PSNR分别达到26.968, 27.880, 28.939 dB, SSIM分别达到0.867, 0.882, 0.947, LPIPS分别达到0.099, 0.141, 0.047。在3个非配对数据集上BRISQUE分别达到25.67, 20.51, 18.80, NIQE值分别达到3.79, 4.09, 4.02。本文为低光照图像增强提供了一种可行的频域-空域联合增强算法。

关键词:图像增强;低光照图像;傅里叶频域;交叉注意力;Transformer

中图分类号:TP394.1;TH691.9 **文献标识码:**A

doi:10.37188/OPE.20253313.2108 **CSTR:**32169.14.OPE.20253313.2108

Three-stage low-light image enhancement based on illumination guidance

ZHANG Yihang, ZHONG Han*

(College of Information and Network Safety, People's Public Security University of China,
Beijing 100038, China)

* Corresponding author, E-mail: zhonghan@ppsuc.edu.cn

Abstract: Addressing challenges such as severe noise, color deviation, and artifacts in low-light imaging, a novel three-stage low-light image enhancement algorithm based on illumination guidance (IG-TSNet) is proposed. This algorithm synergistically integrates the Fourier domain's capability to capture global image information with the Transformer's strength in modeling long-range dependencies within the spatial domain, utilizing illumination guidance to ensure coherent enhancement. IG-TSNet comprises three sequential stages. In the pixel-wise enhancement stage, an adaptive parameter adjustment mechanism is introduced to improve the global representation of the image. During the Fourier reconstruction stage, illumination priors are employed to optimize both amplitude and phase spectra across two channels following Fourier

收稿日期:2025-03-24;修订日期:2025-05-18.

基金项目:中国人民公安大学“双一流”创新研究项目(No. 2023SYL07);高等学校学科创新引智基地项目(No. B20087)

er transformation, enabling comprehensive global image reconstruction. In the cross-attention fusion stage, a lightweight dual-path U-shaped network, incorporating a cross-attention fusion module, is designed to dynamically align Fourier-reconstructed features with illumination-guidance maps. The proposed IG-TSNet was rigorously evaluated on six benchmark datasets for low-light image enhancement, demonstrating superior performance. Qualitative results confirm that the method effectively enhances underexposed regions, suppresses noise without introducing artifacts or patchiness, and preserves color fidelity robustly. Quantitative assessments reveal that IG-TSNet achieves state-of-the-art results across nine evaluation metrics. On three paired datasets, PSNR values of 26.968 dB, 27.880 dB, and 28.939 dB; SSIM values of 0.867, 0.882, and 0.947; and LPIPS values of 0.099, 0.141, and 0.047 were attained, respectively. On three unpaired datasets, BRISQUE scores of 25.67, 20.51, and 18.80 and NIQE values of 3.79, 4.09, and 4.02 were achieved, respectively. This study offers a viable frequency-spatial joint enhancement framework, advancing the field of low-light image enhancement.

Key words: image enhancement; low-light image; Fourier frequency domain; cross-attention; transformer

1 引言

低光照条件是捕获图像数据时客观存在且较为常见的现实场景,在此情况下拍摄的图像不仅影响人眼观感,而且不利于其内部信息的准确获取和后续使用。一般的应对方法是通过调整拍摄设备参数以匹配环境光照影响,如提高感光度,增加曝光时间,加大光圈等,但这并不能完全解决低光照条件下出现的捕获图像噪点严重、色彩偏差、伪影失焦等问题。传感器等硬件的极限、成本与功耗限制使得用数字图像处理技术进行低光照图像增强(Low-Light Image Enhancement, LLIE)成为目前主流的解决方案,其目的是通过算法改善低光照图像的低质量问题,提高人眼视觉感知或服务高层视觉任务。

直方图均衡化^[1-2]和伽马矫正^[3-5]等方法直接调整整幅图像的亮度和对比度。这些方法几乎没有考虑光照因素和图像本身的结构信息,导致增强后的图像在视觉上与真实的正常光照场景不一致。传统方法考虑了光照因素,它通常依赖于Retinex理论^[6]。该理论提出观察到的颜色是场景光照条件(入射光)和物体表面反射性质共同作用的结果。基于该理论,Wang等^[7]设计了一种处理光照不均匀图像时保持自然度和信息的方法,Guo等^[8]提出通过解构先验优化初始估计的光照图。然而,这类方法依赖于人工设计的先验,通常需要繁琐的参数调整,泛化能

力较差。

随着深度学习的快速发展,卷积神经网络(Convolutional Neural Network, CNN)应用于解决LLIE问题。第一类方法直接用CNN来学习从低光照图像到正常光照图像的暴力映射函数,通过设计调整损失函数改善学习效果。但一张图像蕴含的深层次信息(色彩空间、对比度、曝光值等)与各通道都表现出强依赖关系,需要复杂网络结构或更大参数学习且不一定能产生理想效果。第二类方法受Retinex理论的启发,这类方法通常采用多级训练,分别使用不同的CNN对彩色图像进行分解,对反射图进行去噪及光照调整。这些CNN首先独立训练,然后连接在一起进行端到端微调,训练过程繁琐耗时。Wei等^[9]开发了基于Retinex的LLIE网络RetinexNet。Zhang等^[10]提出了KinD方法,该网络与RetinexNet相似,但核心在于引入了更多的训练损失约束。另外还有一类是无监督方法。Guo等^[11]将图像增强任务重新定义为图像特定曲线估计问题,构建了一个轻量级网络ZeroDCE估计像素级和高阶曲线。Jiang等^[12]首次将生成对抗网络(Generative Adversarial Network, GAN)应用于LLIE任务。

计算机视觉领域,CNN具有归纳偏置的先验优势。而在LLIE领域,CNN在捕捉图像各区域亮暗关系、高低频信号分布、颜色亮度联系等全局信息时却表现出局限性。Transformer在获取图像复原任务中关键的长距离依赖和非局部

自相似性上表现出优势。Xu等^[13]利用信号噪声感知 Transformer 与 CNN 开发的混合模型 SNR-Aware, 通过空间变化操作动态增强像素获得优秀的感知质量。Cai等^[14]结合 Retinex 理论, 并且考虑隐藏在低光照图像中的退化因素和点亮过程中引入的噪声, 开发了单阶段训练网络 Retinexformer。尽管上述方法在理论建模或空域信息利用方面取得了显著成果, 但却较少关注对 LLIE 有用的傅里叶频域信息。

目前, 使用傅里叶频域信息进行 LLIE 的探索还处于初期阶段。Yu等^[15]同时探索频率域和空间域的信息来处理图像去雾问题。Huang等^[16]通过傅里叶变换分析不同曝光图像的频率特性, 发现幅度分量包含大部分光照信息, 而相位分量与结构信息相关。Wang等^[17]研究表明, 通过傅里叶分析可以在不增加模型参数的情况下实现全局特征提取, 低光图像的亮度退化主要存在于傅里叶空间中的幅度分量, 而相位分量包含与亮度无关的信息。尽管利用频域信息进行图像增强任务取得了一定进展, 但孤立操纵振幅或相位分量显示出局限性。针对上述问题, 本文提出了一种光照引导的三阶段低光照图像增强算法 IG-TSNet。在第一阶段设计了自适应像素增强模块 (Adaptive Pixel-wise Enhancement Module, APEM), 通过 Transformer 风格的网络结构和内嵌的可学习调整参数对输入的低质量源数据进行像素级自适应调整, 增强图像的深层信息表征。在第二阶段对图像进行傅里叶重建。为了解决傅里叶空间中改变幅度分量进行亮度增强导致的潜在溢出问题和相位分量包含有大量图像结构信息的问题, 使用图像光照信息引导幅度和相位两分量重建。光照通道包含有图像

纹理细节和各区域亮暗关系, 设计了一个 Y 通道增强模块 (Y-channel Enhancement Module, YEM), 使用内嵌的多头注意力机制提取光照引导信息, 精确指导幅度放大和相位调整, 进一步指导第三阶段对空间结构和图像纹理进行重建。在第三阶段, 设计了一个双路径 U 型 CNN-Transformer 混合结构的重建网络, 内嵌的轻量级交叉注意力模块 (Lightweight Cross-Attention, LCA) 将光照信息和傅里叶重建图像通过交叉注意力机制进一步融合。IG-TSNet 不仅有效地增强了图像的亮度, 而且拥有较好的色准表现, 在图像纹理和结构细节方面展现出显著优势。

2 原 理

2.1 网络总体结构

图 1 展示了本文算法 IG-TSNet 的整体网络架构。该方法分为三个阶段: 像素增强阶段, 傅里叶重建阶段和交叉注意力融合阶段。低光照图像送入 IG-TSNet 后首先进入像素增强阶段, 该阶段内部有乘法分支和加法分支, 用于生成像素增强的乘法图和加法图。经过一阶段增强后的输出图像 I_{s1} 随后输入傅里叶重建阶段, 这一阶段使用傅里叶频域信息对图像全局进行重建。该阶段内部设计了亮度分支获取傅里叶重建所需的幅度和相位增强信息, 并且该分支生成的光照引导图将输入交叉注意力融合阶段进行空间细节增强。第三阶段设计了一个内嵌 LCA 的双路径 U 型结构网络, 对经过傅里叶重建的输出 I_{s2} 与光照分支获得的光照引导信息进行交叉注意力融合。经过该阶段增强的结果与第一、二阶段的输出残差相加生成最终的增强结果 I_{s3} 。

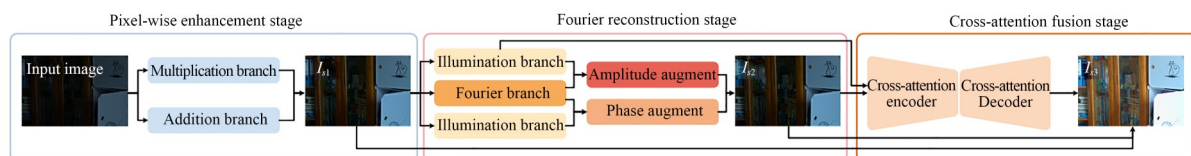


图 1 光照引导的三阶段低光照图像增强网络总体架构

Fig. 1 Overall network architecture of IG-TSNet

2.2 像素增强阶段

由于相机图像处理流程中的非线性处理环

节, 传统的图像增强方法难以有效校正 sRGB 图像中因白平衡错误导致的色彩偏差问题^[18]。因

此,为了更加准确进行图像编辑,经过渲染的 sRGB 图像应该进行逆向处理,以恢复出接近原始传感器数据的 raw-RGB 表示,其抽象过程描述如下:

$$I_i = G(F(I_i)), \quad (1)$$

其中: F 为从 sRGB 到 raw-RGB 的未知重建函数, G 为从 raw-RGB 到 sRGB 的未知相机渲染函数。考虑到相机图像处理链的非线性特性, G 可视为由多个非线性子函数组成的复合函数,每个子函数负责处理特定操作阶段的中间表示。进一步用编解码器架构来建模上述流程,编码器 f 用于模拟逆向映射函数 F ,解码器 g 用于模拟正向渲染函数 G ,且这两部分相互独立工作。将以上建模流程转化成具体的编解码器结构,可表述如下:

$$\hat{I}_i = g_i(f(I_i)). \quad (2)$$

受 Bread^[19]的启发,对于过程 F 采用逐像素最小二乘模型 $f(I_i) = I_i \odot M + A$ 进行模拟。 f 由两个分支组成,分别处理像素级乘法图 M 和加法图 A ,它们具有与输入图像相同的维度,对输入图像进行像素级乘性和加性调整。

如图 2 所示,像素增强阶段首先通过深度可分离卷积层对输入图像进行通道维度的扩展。随后,扩展后的特征图被并行输入到两个结构相同但参数独立的分支中。每个分支包含 3 个级联的 APEM 模块,模块间通过逐元素相加的方式连接。APEM 模块在内部和外部均有残差连接,以缓解堆叠设计产生的特征退化问题。

APEM 内部首先通过深度可分离卷积实现跨通道特征融合,然后采用逐点卷积叠加深度可分离卷积层增强局部像素细节,最后通过 2 个 1×1 卷积层对增强后的像素信息进行充分融合。该模块的创新在于设计的自适应归一化层取代常见的层归一化或批归一化,其内部包含两个可学习光调整参数 α, β 和一个可学习色彩调整矩阵,这些参数在训练过程中能够根据数据集内图像像素分布特点自适应地学习不同光照条件下的最优调整策略,能显著提升模型的环境适应能力。经过双分支处理后,网络通过 3×3 卷积层进行通道降维,并分别使用 ReLU 和 Tanh 激活函数生成用于像素级调整的乘法特征图和加法特征图。

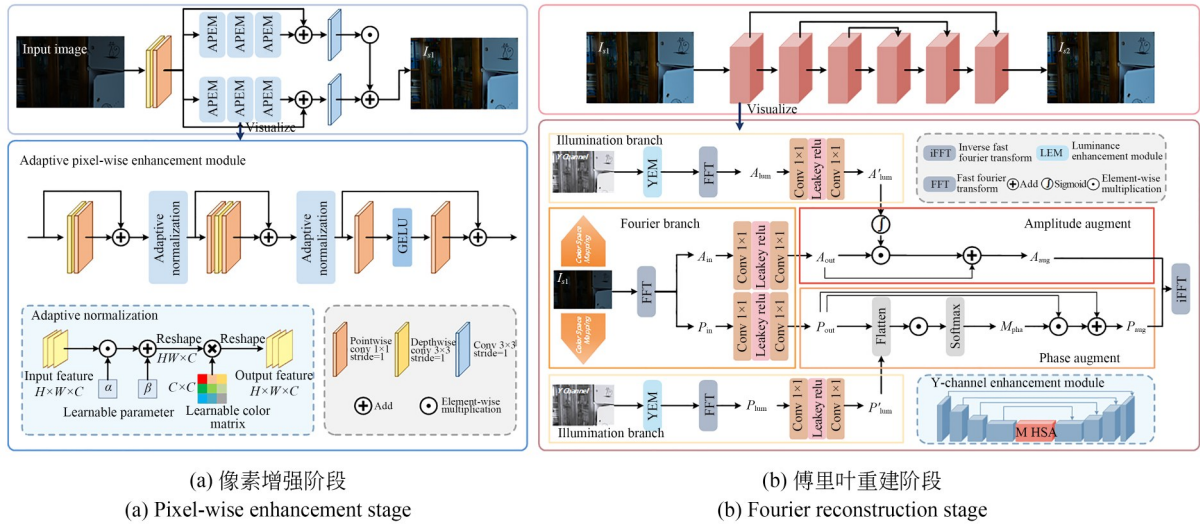


图 2 像素增强阶段与傅里叶重建阶段的网络结构

Fig. 2 Network structure of pixel-wise enhancement and Fourier reconstruction stages

2.3 傅里叶频域信息

为详细拆解傅里叶重建阶段的算法设计细节,这里对数字图像的频域信息进行介绍。傅里叶频域信息是指使用二维离散傅里叶变换

(Two-Dimensional Discrete Fourier Transform, 2D-DFT)将输入图像从空间域转换到频率域得到的频率成分。输入图像的形状为 $H \times W$, H, W 分别表示图像的高和宽。2D-DFT 及其逆变换

公式表示为:

$$F(u, v) = \sum_{h=0}^{H-1} \sum_{w=0}^{W-1} f(x, y) e^{-j2\pi\left(\frac{h}{H}u + \frac{w}{W}v\right)}, \quad (3)$$

$$f(x, y) = F^{-1}(u, v) = \frac{1}{HW} \sum_{u=0}^{H-1} \sum_{v=0}^{W-1} F(u, v) e^{-j2\pi\left(\frac{h}{H}u + \frac{w}{W}v\right)}, \quad (4)$$

其中: x, y 是空域坐标, u, v 是频域坐标, j 为虚数单位, 其逆过程记为 F^{-1} 。式(4)表示在 (u, v) 处的 DFT 值是通过对所有空间点的像素强度乘以一个复数指数函数并进行累加得到的。 F 由复数组成, 可以表示为:

$$F(u, v) = R(F(u, v)) + jI(F(u, v)), \quad (5)$$

其中 R 和 I 分别是实部和虚部。每个复分量可以由幅度分量和相位分量表示。这两个分量分别为:

$$A(F(u, v)) = \sqrt{R^2(F(u, v)) + I^2(F(u, v))}, \quad (6)$$

$$P(F(u, v)) = \arctan\left[\frac{I(F(u, v))}{R(F(u, v))}\right], \quad (7)$$

其中 R 和 I 也可以由下式得到:

$$R(F(u, v)) = A(F(u, v)) \times \cos(P(F(u, v))), \quad (8)$$

$$I(F(u, v)) = A(F(u, v)) \times \sin(P(F(u, v))). \quad (9)$$

2.4 傅里叶重建阶段

该阶段由 6 个具有跳跃连接的相同组件组成, 组件内部操作细节已可视化如图 2(b) 所示。

2.4.1 亮度分支

由于自然场景中往往存在不均匀的光照分布, 同一图像的不同区域通常表现出显著的光照差异。具体而言, 暗区不仅存在严重的噪声干扰和细节丢失问题, 其视觉信息仍然被压缩在有限的动态范围内, 这使得传统的全局增强方法难以实现令人满意的恢复效果。相比之下, 光照充足区域通常具有更丰富的纹理细节和更低的噪声水平, 能够为暗区重建提供有价值的语义上下文信息。因此, 针对 LLIE 任务的特点, 本文提出了一种光照引导的自适应增强策略。

在各种色彩空间中, YCbCr 色彩空间中的亮度分量 Y 表现最好, 受噪声影响最小^[20]。因此, 本文基于 YCbCr 色彩空间中的亮度分量 Y , 通过 YEM 得到光照引导信息。YEM 使用经典的 U

型网络结构, 并在瓶颈层集成多头注意力机制。该模块的下采样过程由 1 个步长为 1 的 conv3×3 层和 3 个步长为 2 的 conv3×3 层组成, 用于捕获不同尺度的特征。瓶颈层多头注意力机制使得该模块能够捕获亮度图中各区域的长程依赖关系, 这对于精确指导不同区域的傅里叶重建和全局交叉注意力融合十分重要。随后, 通过上采样层与跳过连接重建光照引导信息, 使其恢复至原始维度参与傅里叶重建和后续交叉注意力信息融合。受限于全局多头注意力的计算效率问题, 这里只在瓶颈层使用全局注意力。此外, 使用基于插值的上采样代替转置卷积是为了减少该模块的参数量。YEM 处理后获得的光照信息进一步经过傅里叶变换和信息提取单元的处理, 形成可以作用于傅里叶域的引导信息 A'_{lum} 和 P'_{lum} 。

2.4.2 傅里叶分支

首先对输入图像进行傅里叶变换, 将其分解为振幅分量 A_{in} 和相位分量 P_{in} 。振幅分量与图像的整体亮度和对比度相关, 而相位分量则编码了图像的结构信息, 该分支分别优化二者可以更准确控制不同频域特征的增强效果。为了实现自适应地学习频域特征分布, 增强关键频率成分, 同时抑制噪声或冗余分量, 在该阶段设计了一个由 1×1 卷积层和 LeakyRelu 激活函数构成的信息提取模块, 分别对振幅和相位分量进行特征提取。

为了解决在傅里叶域调整幅度分量增强不同光照条件区域时可能导致的过度增强或欠增强问题, 本研究设计了一种基于光照引导的亮度增强操作。该操作通过引入光照先验信息, 实现了幅度分量的自适应调整, 使得幅度增强能够根据图像内部各区域的上下文信息进行精确调控。具体来说, A'_{lum} 经过 sigmoid 操作, 然后逐元素乘以增强图像的傅里叶变换后的幅度分量 A_{out} , 通过残差连接得到增强后的振幅分量 A_{aug} 。其过程可以描述为:

$$A_{aug} = A_{out} \times \text{sigmoid}(A'_{lum}) + A_{out}. \quad (10)$$

傅里叶空间中的相位分量对图像的结构和内容至关重要。由于光照引导信息在亮度分支中已有效解耦了色彩干扰, 因此, 可利用其内部保留的纹理细节等关键结构信息引导相位分量

增强。受转置注意力^[21]的启发,本文设计了一种基于光照引导的相位增强操作。该操作通过在傅里叶空间中对光照引导信息和增强图像的相位分量特征进行融合,得到相位增强引导信息 M_{pha} 。相位增强信息的计算过程如下:

$$M_{pha} = \text{soft max}(\text{flatten}(P_{out}) \times \text{flatten}(P'_{lum})). \quad (11)$$

然后 P_{out} 与 M_{pha} 相乘并通过残差连接相加,得到增强后的相位分量 P_{aug} ,即:

$$P_{aug} = P_{out} \times M_{pha} + P_{out}. \quad (12)$$

经过以上流程处理,得到增强的幅度分量和相位分量,然后通过傅里叶逆变换回到空间域,从而得到第二阶段傅里叶重建后的图像,即:

$$I_{s1} = F^{-1}(A_{aug} \times \cos(P_{aug}) + A_{aug} \times \sin(P_{aug})). \quad (13)$$

2.5 交叉注意力信息融合阶段

傅里叶空间在捕捉图像全局信息表现出优势,但是缺乏对空间细节的增强,在傅里叶变换和特征传递过程中可能会引入额外的噪声和特征损失。为此,本文基于 CNN-Transformer^[22]混合架构设计了交叉注意力融合阶段,进一步细化

图像细节和纹理结构。

如图 3 所示,该阶段由一个双路径 U 型结构的网络进行处理,一路输入为第二阶段增强输出 I_{s2} ,一路输入为第二阶段光照引导信息 I_{lum} 。该网络包含三次下采样和三次上采样操作。每次下采样将高和宽减半并将通道数加倍,每次上采样将高和宽加倍,通道数减半。6 个 LCA 模块嵌入到每次下采样和上采样操作后,进一步融合两路特征。为了进一步降低 Transformer 的计算复杂度,本文设计了一个具有线性复杂度的 LCA 模块,如图 3(a)所示。

如图 3(b)所示,首先,对于输入特征图 I_{s2}, I_{lum} 使用 $\text{conv}3 \times 3$ 深度卷积 W_d 和 $\text{conv}1 \times 1$ 逐点卷积 W_p 叠加组成的深度可分离卷积操作,获取计算注意力所需的 query, key 和 value,并且 key 和 value 由同一卷积操作后沿通道维度对半分离所得。该操作取代线性投射层,并把多头操作从常规的作用于图像宽高尺度上改为作用于通道维度上,大幅降低推理时间和计算复杂度。计算过程可表示为:

$$Q_{s2} = W_p^Q W_d^Q I_{s2}, \quad (14)$$

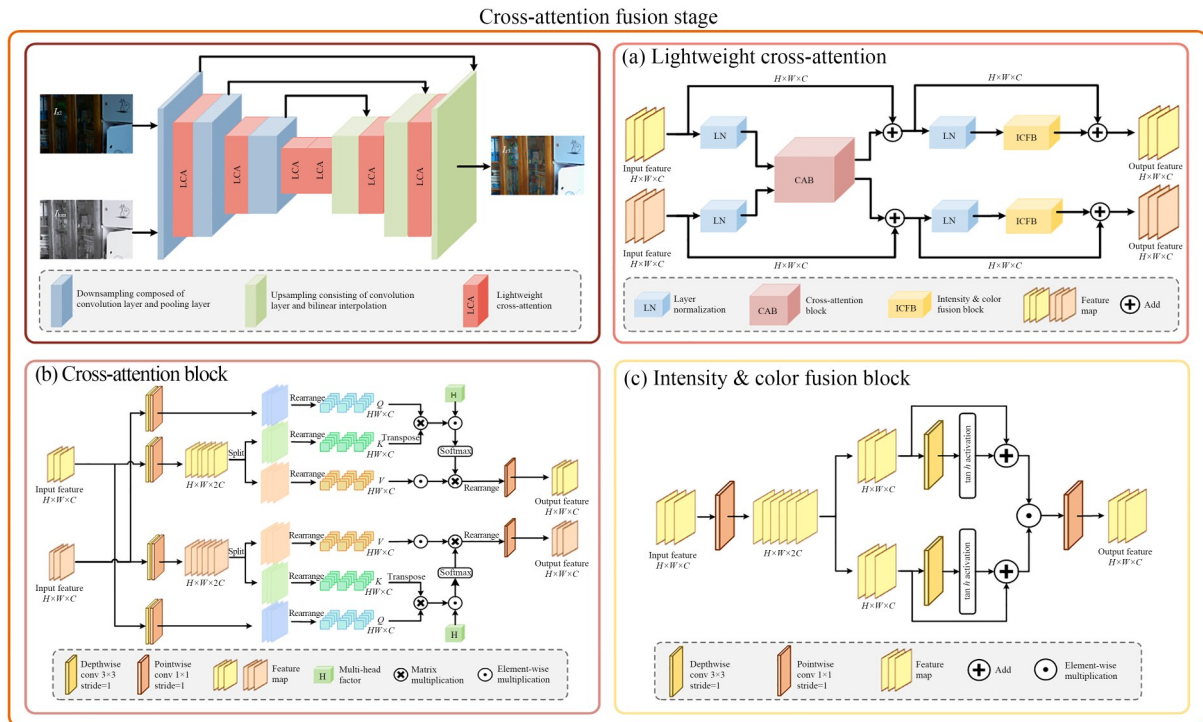


图 3 交叉注意力融合阶段的网络结构

Fig. 3 Network structure of cross-attention fusion stage

$$K_{\text{lum}} V_{\text{lum}} = W_p^{KV} W_d^{KV} I_{\text{lum}}, \quad (15)$$

$$K_{\text{lum}}, V_{\text{lum}} = \text{chunk}(K_{\text{lum}} V_{\text{lum}}, 2, \text{dim} = 1). \quad (16)$$

接着将 Q, K, V 沿着通道维度分为 k 个头可得到多头注意力操作所需 Q_i, K_i, V_i :

$$Q_{s2} = [Q_{s2}^1, Q_{s2}^2, \dots, Q_{s2}^k], \quad (17)$$

$$K_{\text{lum}} = [K_{\text{lum}}^1, K_{\text{lum}}^2, \dots, K_{\text{lum}}^k], \quad (18)$$

$$Q_{\text{lum}} = [Q_{\text{lum}}^1, Q_{\text{lum}}^2, \dots, Q_{\text{lum}}^k]. \quad (19)$$

最后,交叉注意力的计算公式如下:

$$\hat{I}_{s2} = W_p I_{s2} \otimes \text{soft max} \left(\frac{Q_{s2} K_{\text{lum}}^T}{\alpha_H} \right) + I_{s2}, \quad (20)$$

其中: α_H 与 SA^[23] 类似是一个可学习的多头因子,自适应地调节矩阵乘法; $\otimes$ 代表矩阵乘法。以上各层都是无偏置的;输出特征 $\hat{I}_{s2} \in \mathbb{R}^{C \times H \times W}$, 为多头交叉注意力计算后的 I_{s2} 与其残差之和。另一路的计算过程相同,只需改变输入。

受 Transformer 中前馈层的启发,上述两路经过 CAB 处理后的特征张量经过层归一化后, (Intensity & Color Fusion Block, ICFB) 进一步融合双路径特征,增加模型非线性能力,如图 3(c) 所示。ICFB 先使用 1×1 深度卷积将输入特征张量通道数加倍,再沿通道维度对半分离为两支路,紧接 3×3 深度卷积和 tanh 激活函数,采用残差连接保留激活前特征,随后分离张量经过元素级乘法融合和输出映射层得到融合后的张量输出。其过程如下:

$$F_{\text{out}} = W_p ((\tan h(W_d F_{\text{in}}^1) + F_{\text{in}}^1) \odot (\tanh(W_d F_{\text{in}}^2) + F_{\text{in}}^2)) + F_{\text{in}}, \quad (21)$$

其中: $\odot$ 代表元素级乘法, $F_{\text{in}}^1 = W_p^1 W_d^1 F_{\text{in}}, F_{\text{in}}^2 = W_p^2 W_d^2 F_{\text{in}}$ 且 $F_{\text{in}}^1, F_{\text{in}}^2 \in \mathbb{R}^{C \times H \times W}$ 。

2.6 损失函数

IG-TSNet 采用端到端的训练方式。为了在增强图像时进一步考虑图像的结构信息、感知信息和边缘细节,本文设计了一种新的混合损失函数,包含 L1 损失函数 L_1 , 结构相似性指数损失函数 L_{ssim} ^[24], 感知损失函数 L_{perc} ^[25] 和边缘损失函数 L_{edge} ^[26]。其公式为:

$$L = \omega_1 L_1(\hat{x}, x) + \omega_{\text{ssim}} L_{\text{ssim}}(\hat{x}, x) + \omega_{\text{perc}} L_{\text{perc}}(\hat{x}, x) + \omega_{\text{edge}} L_{\text{edge}}(\hat{x}, x), \quad (22)$$

式中: $\omega_1, \omega_{\text{ssim}}, \omega_{\text{perc}}, \omega_{\text{edge}}$ 均为实验中用于权衡损失函数 L 的权重系数。

3 实验及结果分析

3.1 数据集与实验设置

本文使用 3 个 LLIE 配对基准数据集对所提方法进行评估,包括 LOLv1^[9], LOLv2-Real^[27], LOLv2-Syn^[27]。LOLv1 数据集由 500 对低光/正常光照图像对组成,所有图像均在真实低光环境下通过专业相机采集,并经过人工调整曝光和感光度得到对应的正常光版本,因此包含了真实情况下的退化特征和属性。LOLv2-Real 数据集由 789 对低光/正常光照图像对组成,是对 LOLv1 数据集的扩充,覆盖了室内外、夜间、背光等复杂场景,用于评价算法在真实场景下的泛化能力。LOLv2-Syn 数据集由 1 000 对图像组成,其低光版本是通过分析多个数据集的统计特性后,采用专业图像处理软件对正常光图像进行模拟退化得到,用于评估算法在不同场景和物体下的稳定性。训练集和测试集在 3 个数据集上分别按照 485:15, 689:100 和 900:100 的比例进行划分。实验使用 3 个非配对基准数据集 DICM^[28], MEF^[29], NPE^[7] 来评估本文算法在真实低光照场景中的泛化能力,它们具有涵盖场景丰富,包含多分辨率图像的特点。

实验基于 PyTorch 2.4.0 深度学习框架,在 NVIDIA GTX TITAN GPU(CUDA 12.1)平台上进行单卡训练与测试。模型优化采用 Adam 优化器 ($\beta_1=0.9, \beta_2=0.999$)^[30], 初始学习率设为 1×10^{-4} , batch size 设置为 8。训练过程持续至少 500 个 epoch,并采用余弦退火方案^[31]实现学习率的稳定衰减。实验中,还将训练数据随机剪裁为 256×256 的图像块,并采用水平和垂直翻转的方式对训练数据进行扩充。

3.2 评价指标

对于配对数据集,本文采用峰值信噪比 (Peak Signal-to-Noise Ratio, PSNR) 和结构相似性指数 (Structural Similarity, SSIM)^[24] 作为重建图像的失真度量。为了更加客观评估恢复图像的感知质量,在上述两个常见指标的基础上增加由 AlexNet^[32] 计算的图像感知相似度 (Learned Perceptual Image Patch Similarity, LPIPS) 分数作为感知度量指标。对于非配对数据集,本文采用 BRISQUE^[33] (Blind/Referenceless Image Spatial

Quality Evaluator) 和 NIQE^[34] (Natural Image Quality Evaluator)来评估图像质量。

3.3 主观评价

为验证模型在不同场景下的泛化能力,将本文算法与 KinD^[10], MIRNet^[35], RUAS^[36], LLFlow^[37], Restormer^[21], LLFormer^[38], SNR-Aware^[13]和 Retinexformer^[14]8种有代表性的算法进行真实场景下室内外低光照图像增强结果对比。这里从 LOLv2-Real 数据集中分别选取了一幅室外低光环境图像和一幅室内低光环境图像作为比较对象,所展示结果图均是通过公开可用的代码在推荐设置下实验所得。

不均匀光照形成的大光比低光场景是室外

环境中增强难度较大的场景,需要算法具备对图像不同光照条件区域采取不同增强策略的能力。图4展示了该场景下本文方法与其他算法的比较结果。KinD, RUAS的增强结果存在明显色偏和细节丢失问题,与人眼自然观感差距较大;MIRNet, LLFlow对暗区的亮度提升不足,同时远处天空区域因过度增强而出现过曝现象,导致楼房边缘信息丢失;Restormer画面整体出现偏色问题;SNR-Aware的在后方楼房区域出现轻微过曝;Retinexformer的整体亮度略暗,特别是汽车后方的路面部分存在较大面积的暗区。相比之下,LLFormer与本文算法在该场景下展现出更优的增强效果。

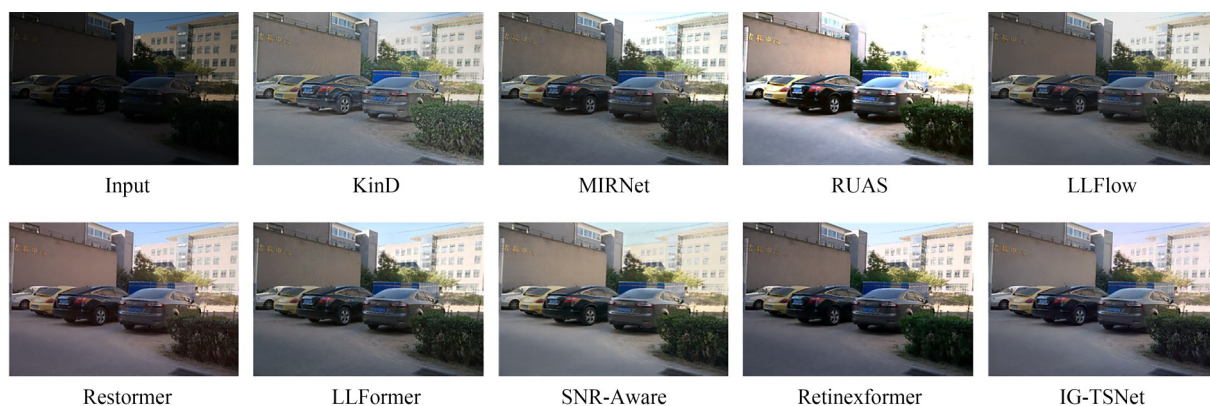


图4 室外低光环境图像增强结果对比

Fig. 4 Comparison of image enhancement results in outdoor low-light environments

在无光环境下,由于严重的色彩和结构信息丢失,室内场景的图像增强面临巨大挑战,对算法的色彩还原和细节恢复能力提出了极高的要求。实验对比了各算法在该极端场景下的增强效果,如图5所示,不同算法性能差异显著:KinD和RUAS延续了其在室外场景中的增强风格,仍然存在明显的色彩偏差问题;MIRNet, LLFlow, Restormer和Retinexformer普遍存在亮度提升不足的问题,导致增强后的图像整体偏暗;LLFormer虽然实现了显著的亮度提升,但地板区域出现了与真实色彩不符的泛白现象;SNR-Aware则在地面区域表现出色彩不均匀的缺陷。相比之下,本文算法不仅实现了显著的亮度提升,而且在色准方面表现出色,更好地还原了场景的真实视觉效果。

为进一步评估本文算法的增强结果在视觉

观感和细节保留方面的突出优势,在 LOLv1 和 LOLv2-Real 和 LOLv2-Syn 3个配对数据集上进行了更加细致的对比实验。在 LOLv1 数据集上,本文选取了11种代表性算法(RetinexNet^[9], KinD^[10], ZeroDCE^[11], MIRNet^[35], RUAS^[36], LLFlow^[37], EnlightenGAN^[12], Restormer^[21], SNR-Aware^[13], LLFormer^[38]和 Retinexformer^[14])的增强结果进行对比分析;在 LOLv2-Real 和 LOLv2-Syn 数据集上,则重点比较了 KinD, MIRNet, RUAS, LLFlow, Restormer, LLFormer, SNR-Aware, Retinexformer 8种算法的增强效果。所有结果均采用以下两种方式获取:(1)直接采用算法发布者公开的增强结果;(2)基于算法开源代码,在推荐参数设置下重新生成。对所有对比结果均进行了局部放大处理,以呈现最佳的视觉对比效果。

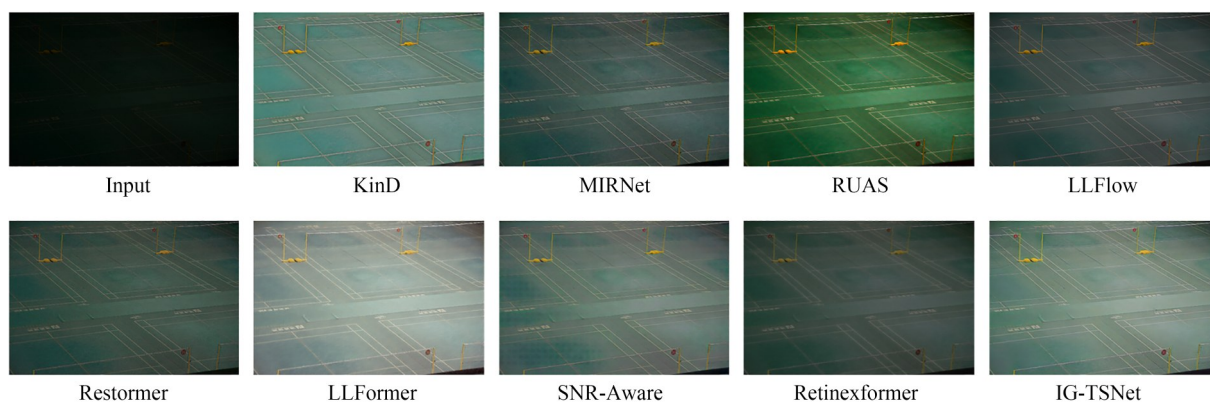


图5 室内低光环境图像增强结果对比

Fig. 5 Comparison of image enhancement results in indoor low-light environments

图6展示了11种算法在LOLv1数据集上的对比结果(彩图见期刊电子版)。RetinexNet的增强结果不仅产生大量噪声干扰,还出现严重的色彩失真现象;KinD, RUAS和ZeroDCE在上下两幅测试图像中均表现出明显的色偏问题,特别是在下图中存在大面积亮度提升失败区域,红绿两框标注的局部放大区域几乎无法辨识内容细节;EnlightenGAN的下图均出现明显色偏,并且绿框标注区域墙面图案涂抹感严重;MIRNet,

Restormer, SNR-Aware在下图红框标注位置条状纹理丢失严重不能辨认,绿框标注位置墙面图案涂抹感严重;LLFormer和Retinexformer上图红框放大区域橱窗玻璃纹理涂抹感严重,下图红框位置条状纹理不够清晰。相比之下,本文算法在两幅图像的色准和细节保留方面优势明显,具体表现在橱窗玻璃纹理锐度更高,条状纹理和墙面图案清晰。

图7展示了8种算法在LOLv2-Real数据集

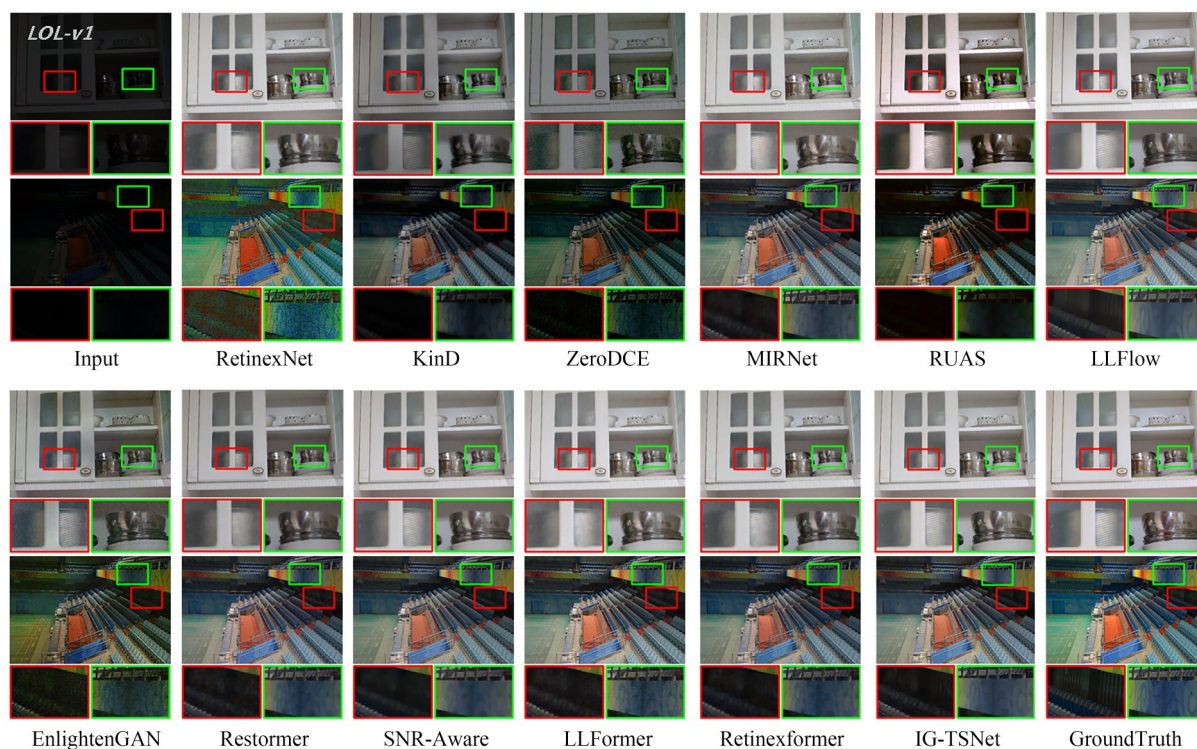


图6 LOLv1数据集不同算法的增强结果细节对比

Fig. 6 Detail comparison of enhancement results by various algorithms on LOLv1 dataset

上的对比结果。除了RUAS以外,其余算法都表现出较好的整体观感。放大凳脚细节可见KinD, MIRNet, Restormer, LLFormer, Retinexformer均在增强后产生较多噪点,特别是位于绿框极低光区域难以保持画面纯净;相比之下,LLFlow, SNR-Aware和本文算法均能在提升画面

亮度的同时抑制噪声。

图8展示了8种算法在LOLv2-Syn数据集上的对比结果。对于该幅包含高光比和“云团”复杂元素的图像, KinD, MIRNet, RUAS, Retinexformer产生难以接受的增强效果; LLFlow, LLFormer, SNR-Aware画面整体偏暗, 颜色偏黑;

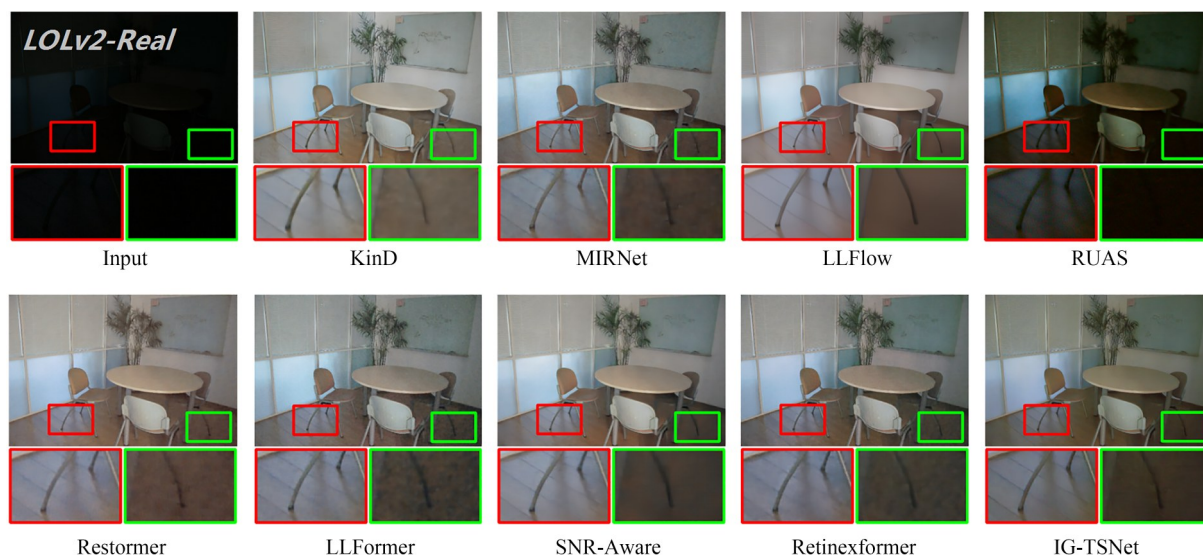


图7 LOLv2-Real数据集不同算法的增强结果细节对比

Fig. 7 Detail comparison of enhancement results by various algorithms on LOLv2-Real dataset

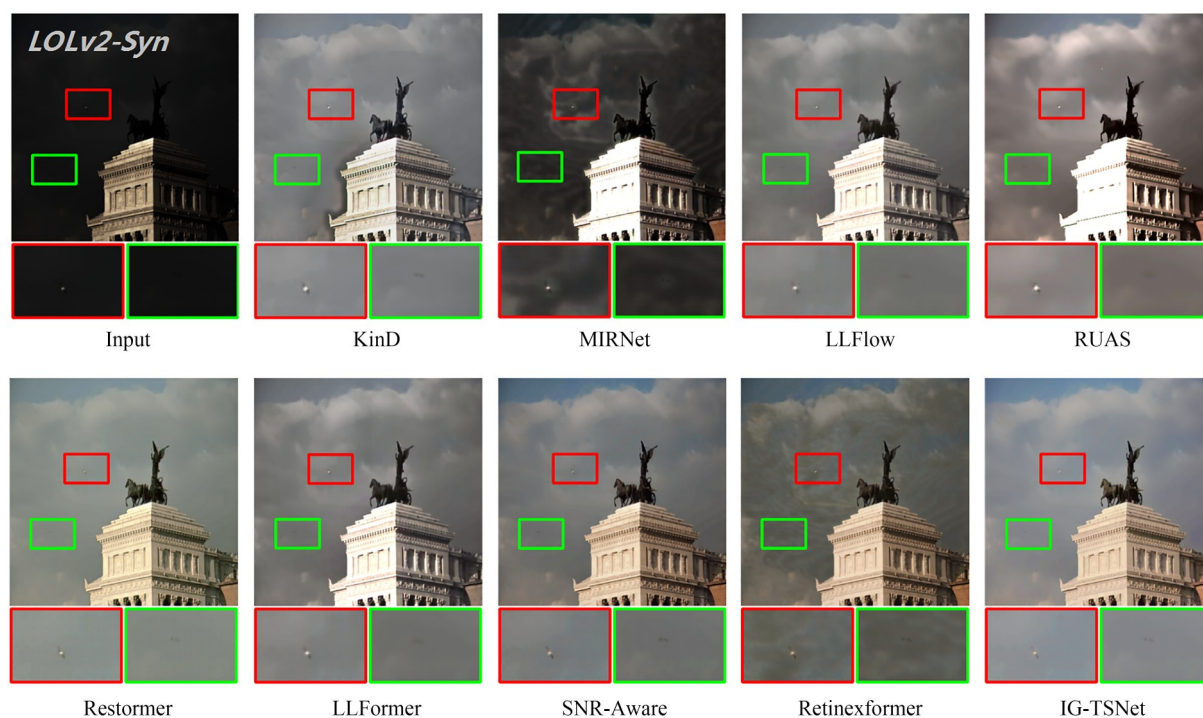


图8 LOLv2-Syn数据集不同算法的增强结果细节对比

Fig. 8 Detail comparison of enhancement results by various algorithms on LOLv2-Syn dataset

Restormer 画面整体泛黄;相比之下,本文算法在画面观感上表现突出,并且通过放大细节可见本文算法在“云团”复杂背景下依然保留了较为清晰的“飞鸟”细节。

为验证本文算法对未训练真实低光照场景数据的泛化能力,在 LOLv2-Real 数据集上训练的模型直接应用于 DICM,MEF,NPE 3 个非配对数据集,并与 MIRNet,SNR-Aware,Retinexformer 3 种算法公开的增强结果进行可视化对比分析。

图 9 对比图像分别选自 3 个数据集,展示了室外高光比、背光、低光照环境下所提算法与其他同类算法的比较结果。在第一行高光比环境图像中,因背景天空与房屋的不均匀光照,容易出现提升亮度的同时造成天空过曝的情况,例如

MIRNet,SNR-Aware 因过曝丢失天空与房顶结构信息。而基于光照引导策略的方法 Retinexformer 和本文算法在提升画面亮度的同时抑制了高光区域的过曝问题。在第二行背光环境图像中,因背光导致前景中信息几乎不可见。SNR-Aware 的结果出现了明显的“网格”效应,Retinexformer 画面亮度 and 对比度不足,MIRNet 与本文算法在该场景下表现较好。在第三行低光照环境下,SNR-Aware 的增强结果出现严重的“网格”效应,相比之下 MIRNet,Retinexformer 和本文算法的增强效果较好。

综上所述,本文算法在 3 个配对数据集和 3 个非配对数据集所包含的各类场景上均表现出稳健且优异的增强效果,在色准保持、细节保留、场景适应性、跨数据集泛化能力等方面具有显著优势。

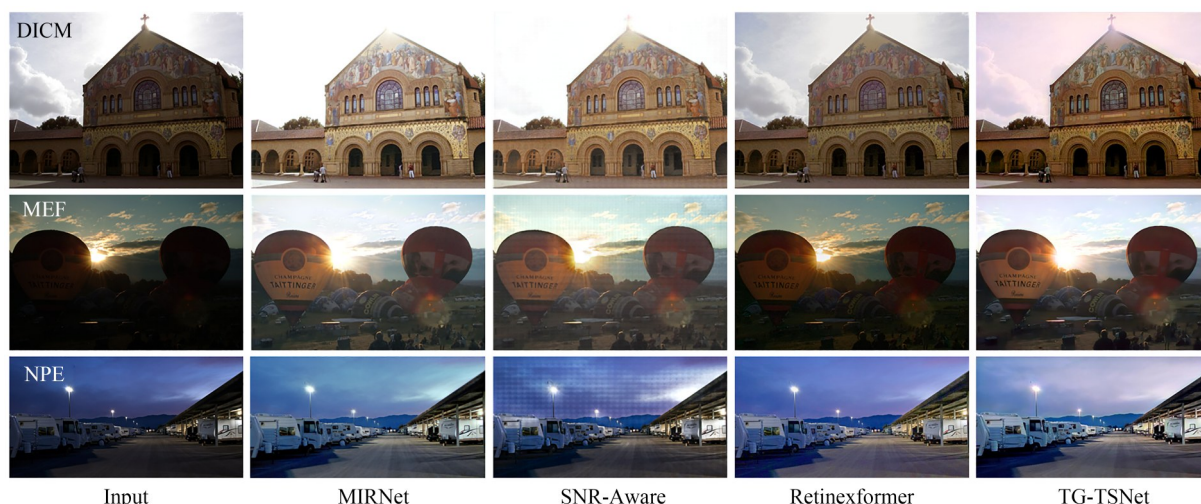


图 9 三个非配对数据集增强结果对比

Fig. 9 Comparison of enhancement results on three unpaired datasets

3.4 客观评价

表 1 对比了 RetinexNet, KinD, ZeroDCE, 3DLUT, DBRN, MIRNet, RUAS, LLFlow, EnlightenGAN, Restormer, LEDNet, SNR-Aware, PairLIE, LLFormer, Retinexformer, FourLLIE 和 PRNet 17 种 LLIE 前沿算法在 3 个配对基准数据集上的 PSNR 和 SSIM 指标。因为并非所有对比算法都涵盖这 3 个数据集上的测试结果,因此表 2 选取其中 11 种算法进行了 LPIPS 指标的对比分析。为便于结果分析,表格中使用粗体标注各项指标的最优结果。Normal 和 GT Mean 分别表

示没有使用 GroundTruth 进行伽马矫正计算的指标和进行伽马矫正后计算的指标。

从评估指标来看,更高的 PSNR 和 SSIM 值,以及更低的 LPIPS 值,通常反映出增强图像具有更好的真实场景还原度和视觉质量。定量结果表明,所提算法在多数情况下都取得了最优或次优结果,并且将模型参数量控制在 1.81M,计算量控制在 14.49G FLOPs。与近期基于傅里叶域的方法 FourLLIE 相比,本算法采用频域振幅相位分离操作机制,并结合频-空域联合增强策略,在 3 个数据集上 PSNR 和 SSIM 指标均有提升,

表 1 不同算法在 LOL 数据集(v1和v2)上的 PSNR/SSIM 对比实验结果

Tab. 1 Comparative experimental results of PSNR/SSIM for different algorithms on LOL datasets (v1 and v2)

Method	Complexity		LOLv1				LOLv2-Real				LOLv2-Syn			
			Normal		GT Mean		Normal		GT Mean		Normal		GT Mean	
	Params/M	FLOPs/G	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
RetinexNet ^[9]	0.84	584.47	16.774	0.419	18.915	0.427	16.097	0.401	18.323	0.447	17.137	0.762	19.099	0.774
KinD ^[10]	8.02	34.99	17.650	0.775	20.860	0.802	14.740	0.641	17.544	0.669	13.290	0.578	16.259	0.591
ZeroDCE ^[11]	0.075	4.83	14.861	0.559	21.880	0.640	16.059	0.580	19.771	0.671	17.712	0.815	21.463	0.848
3DLUT ^[40]	0.59	7.67	14.350	0.445	21.350	0.585	17.590	0.721	20.190	0.745	18.040	0.800	22.173	0.854
DRBN ^[41]	5.47	48.61	16.290	0.617	19.550	0.746	20.290	0.831	—	—	23.220	0.927	—	—
MIRNet ^[35]	31.76	785.10	24.100	0.845	26.519	0.856	20.020	0.820	27.173	0.865	21.940	0.876	25.955	0.898
RUAS ^[36]	0.003	0.83	16.405	0.500	18.654	0.518	15.326	0.488	19.061	0.510	13.765	0.638	16.584	0.719
LLFlow ^[37]	17.42	358.40	21.149	0.854	24.998	0.871	17.433	0.831	25.421	0.877	24.807	0.919	27.961	0.930
EnlightenGAN ^[12]	114.35	61.01	17.480	0.651	20.003	0.691	18.230	0.617	—	—	16.570	0.734	—	—
Restormer ^[21]	26.13	144.25	22.365	0.816	26.682	0.853	18.693	0.834	26.116	0.853	21.413	0.830	25.428	0.859
LEDNet ^[39]	7.07	35.92	20.627	0.823	25.470	0.846	19.938	0.827	27.814	0.870	23.709	0.914	27.367	0.928
SNR-Aware ^[13]	4.01	26.35	24.610	0.842	26.716	0.851	21.480	0.849	27.209	0.871	24.140	0.927	27.787	0.941
PairLIE ^[8]	0.33	20.81	19.510	0.736	23.526	0.755	19.885	0.778	24.025	0.803	21.351	0.857	25.326	0.873
LLFormer ^[38]	24.55	22.52	23.649	0.816	25.758	0.823	20.056	0.792	26.197	0.819	24.038	0.909	28.006	0.927
Retinexformer ^[14]	1.53	15.85	25.153	0.845	27.140	0.850	22.794	0.840	27.694	0.856	25.670	0.930	28.992	0.939
FourLLIE ^[17]	0.12	2.57	22.526	0.831	26.037	0.849	22.127	0.843	27.591	0.871	24.763	0.924	28.121	0.926
PRNet ^[42]	1.93	8.08	23.452	0.837	26.860	0.856	21.314	0.848	27.661	0.883	25.274	0.931	28.657	0.943
Ours	1.81	14.49	23.202	0.849	26.968	0.867	22.260	0.856	27.880	0.882	25.225	0.936	28.939	0.947

表 2 不同算法在 LOL 数据集(v1和v2)上的 LPIPS 对比实验结果

Tab. 2 Comparative experimental results of LPIPS for different algorithms on LOL datasets (v1 and v2)

Method	LOLv1		LOLv2-Real		LOLv2-Syn		Complexity
	Normal	GT Mean	Normal	GT Mean	Normal	GT Mean	FLOPs/G
RetinexNet	0.474	0.470	0.543	0.519	0.255	0.247	587.47
KinD	0.207	0.201	0.375	0.368	0.460	0.435	34.99
MIRNet	0.131	0.128	0.144	0.142	0.058	0.057	758.10
RUAS	0.270	0.257	0.309	0.296	0.364	0.341	0.83
LLFlow	0.119	0.117	0.176	0.158	0.067	0.063	358.40
EnlightenGAN	0.322	0.317	0.309	0.301	0.220	0.213	61.01
Restormer	0.141	0.138	0.231	0.214	0.144	0.139	144.25
SNR-Aware	0.151	0.150	0.156	0.150	0.057	0.056	26.35
LLFormer	0.175	0.167	0.211	0.209	0.066	0.061	22.52
LEDNet	0.118	0.113	0.120	0.114	0.061	0.056	35.92
Retinexformer	0.131	0.129	0.171	0.166	0.059	0.056	15.85
Ours	0.105	0.099	0.154	0.141	0.051	0.047	14.49

Normal 情况下 PSNR 值分别提升 0.676, 0.133, 0.462 dB, SSIM 值分别提升 0.018, 0.013, 0.012; GT Mean 情况下 PSNR 值分别提升 0.931, 0.289, 0.818 dB, SSIM 值分别提升

0.024, 0.011, 0.021。与同样采用光照引导策略的算法 Retinexformer 相比, 本文算法在 FLOPs 降低了 9% 的情况下, 在 3 个数据集上 SSIM 和 LPIPS 指标均有提升, Normal 情况下 SSIM 值分

别提升 0.004,0.013,0.006,LPIPS 值分别提高 20%,9%,13%;GT Mean 情况下 SSIM 值分别提升 0.017,0.026,0.008,LPIPS 值分别提高 23%,15%,16%。与同样使用 Transformer 技术的方法 Restormer, SNR-Aware, LLFormer 和 Retinexformer 相比,本文算法拥有较低模型计算量的同时,在 3 个数据集上的 SSIM 和 LPIPS 指标排名第一。以上数据表明,所提算法在色彩和结构信息上恢复效果更好,对噪声的抑制能力更加出色。综合主观评价结果,本文算法的增强效果更加稳定,对不同场景的适应性更强。

为进一步定量评估本文算法对未训练数据集的泛化能力,将在 LOLv2-Real 数据集上训练的模型直接应用于 DICM,MEF,NPE 3 个非配对数据集,并选取 KinD, ZeroDCE, MIRNet, RUAS, LLFlow, SNR-Aware, PairLIE, Retinexformer 8 种算法进行 BRISQUE 和 NIQE 指标的对比实验。通常更低的 BRISQUE 值表明图像噪声、伪影等失真消除程度更好,更低的 NIQE 值表明图像色彩、纹理真实性更接近视觉自然性。对比结果如表 3 所示,本文算法在 DICM 和 NPE 数据集上两项指标均位于次优位置,在 MEF 数据集上的指标欠佳,但综合主观评价和实验数据,所提算法在消除噪声/伪影和保持自然视觉效果方面具有更出色的表现。

表 3 不同算法在 DICM/MEF/NPE 非配对数据集上的 BRISQUE/NIQE 对比实验结果

Tab. 3 Comparative experimental results of BRISQUE/NIQE for different algorithms on unpaired datasets (DICM/MEF/NPE)

Method	DICM		MEF		NPE	
	BRIS	NIQE	BRIS	NIQE	BRIS	NIQE
KinD	48.72	5.15	49.94	5.47	36.85	4.98
ZeroDCE	27.56	4.58	17.32	4.93	20.72	4.53
MIRNet	28.86	3.98	28.63	3.98	25.56	4.33
RUAS	38.75	5.21	23.68	3.83	47.85	5.53
LLFlow	26.36	4.06	30.27	4.70	28.86	4.67
SNR-Aware	37.35	4.71	31.28	4.18	26.65	4.32
PairLIE	33.31	4.03	27.53	4.06	28.27	4.18
Retinexformer	14.10	3.83	11.78	3.67	18.18	3.94
Ours	25.67	3.79	20.51	4.09	18.80	4.02

3.5 消融实验

为验证本文所提出三阶段网络的每一阶段对 LLIE 任务的有效性,本文在 3 个数据集上进行了消融实验对比,每组消融实验在相同实验环境下进行,实验结果如表 4 所示。针对本文方法中提到的三个阶段设计了五组实验:(1)仅有一阶段自适应像素增强;(2)自适应像素增强结合傅里叶重建;(3)自适应像素增强结合交叉注意力信息融合;(4)傅里叶重建阶段结合交叉注意力信息融合阶段;(5)三阶段光照引导增强网络。

表 4 消融实验定量分析结果

Tab. 4 Quantitative analysis result of ablation experiments

Group	Model			LOL-v1			LOL-v2-Real			LOL-v2-Syn		
	一阶段	二阶段	三阶段	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS
1	✓			18.105	0.736	0.267	18.389	0.772	0.249	19.852	0.873	0.129
2	✓	✓		18.656	0.774	0.252	21.895	0.837	0.183	21.486	0.894	0.164
3	✓		✓	22.370	0.826	0.145	21.430	0.840	0.180	23.710	0.917	0.107
4		✓	✓	22.269	0.824	0.140	21.330	0.842	0.146	23.249	0.915	0.090
5	✓	✓	✓	23.202	0.849	0.105	22.260	0.856	0.154	25.225	0.936	0.051

当仅使用像素增强阶段时,在 3 个数据集上的 PSNR 值均达到 18 dB 以上,甚至超过表 1 中的部分方法;对比第四、五组实验,在增加像素增强阶段后,3 个数据集上的 3 个指标均有明显提高,PSNR 值提升约为 1 dB,表明该阶段在像素

增强上的有效性。对比第一、三组实验,当像素增强结合傅里叶重建时,3 个数据集上的 3 个指标与第一组实验相比均有提高;对比第三、五组实验,像素增强阶段结合交叉注意力融合阶段,在插入傅里叶重建阶段过后,3 个数据集上的 3

个指标均有显著提高,说明本文提出的傅里叶重建阶段对 LLIE 任务的有效性。对比第二、五组实验,增加交叉注意力融合阶段后,三项指标显著提高,对于 LPIPS 涨幅分别达到 59%,35%,69%,说明第三阶段交叉注意力融合对于增强图像视觉感知质量提升明显。三阶段结合后,IG-TSNet 取得最佳结果,说明本文所提三阶段增强网络的每一部分对增强结果都起到关键性作用,三阶段协作达到最优增强性能。

4 结 论

本文提出了一种基于光照引导的三阶段低光照图像增强算法 IG-TSNet。该算法结合了傅里叶变换在全局信息表征方面的优势与 Transformer 在空域长程依赖建模方面的特性,实现了对低光照图像的高质量增强。架构设计上,IG-TSNet 采用三阶段渐进式增强。首先,通过参数自适应机制有效提高低光照源数据的全局表征;然后,从输入图像的亮度通道(Y通道)提取光照先验信息,随后利用这些引导信息对傅里叶域的幅度和相位分量进行解耦优化处理,实现了对幅度信息的自适应增强和相位信息的精准校正;最后,设计

了一个内嵌轻量交叉注意力机制的双路径 U 型网络,有效整合傅里叶重建结果与光照引导信息,实现空间细节的精准增强。所提算法在 6 个 LLIE 基准数据集上进行了评估。实验结果表明,IG-TSNet 可以有效提高欠曝光区域的亮度,可靠抑制噪声不会引入斑块和伪影,并稳健保持图像色准。IG-TSNet 的 9 项指标均达到最优。在 3 个配对数据集上的 PSNR 分别达到 26.968,27.880,28.939 dB,SSIM 分别达到 0.867,0.882,0.947, LPIPS 分别达到 0.099,0.141,0.047。在 3 个非配对数据集上的 BRISQUE 分别达到 25.67,20.51,18.80,NIQE 值分别达到 3.79,4.09,4.02。基于光照引导机制的频域-空域联合增强算法实现了显著的低光照图像增强效果。然而,考虑到实际应用场景对计算效率的要求,未来工作将探索更高效的光照引导方案并优化网络架构,以提高算法在实际应用中的效率。

作者贡献声明:

张易航:方法提出,论文构思和撰写,实验设计及数据整理分析;

钟寒:方法指导,论文审核与编辑,提供资源。

参考文献:

- [1] LIU Y-F, GUO J-M, YU J C. Contrast enhancement using stratified parametric-oriented histogram equalization[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2017, 27(6): 1171-1181.
- [2] MONIRUZZAMAN M, SHAFUZZAMAN M, HOSSAIN M F. Brightness preserving bi-histogram equalization using edge pixels information[C]. *2013 International Conference on Electrical Information and Communication Technology (EICT)*. February 13-15, 2014. *Khulna, Bangladesh*. IEEE, 2014: 1-5.
- [3] HUANG S C, CHENG F C, CHIU Y S. Efficient contrast enhancement using adaptive gamma correction with weighting distribution[J]. *IEEE Transactions on Image Processing*, 2013, 22(3): 1032-1041.
- [4] WANG Z G, LIANG Z H, LIU C L. A real-time image processor with combining dynamic contrast ratio enhancement and inverse gamma correction for PDP[J]. *Displays*, 2009, 30(3): 133-139.
- [5] RAHMAN S, RAHMAN M M, ABDULLAH-AL-WADUD M, et al. An adaptive gamma correction for image enhancement[J]. *EURASIP Journal on Image and Video Processing*, 2016, 2016(1): 35.
- [6] LAND E H. The retinex theory of color vision[J]. *Scientific American*, 1977, 237(6): 108-128.
- [7] WANG S H, ZHENG J, HU H M, et al. Naturalness preserved enhancement algorithm for non-uniform illumination images[J]. *IEEE Transactions on Image Processing*, 2013, 22(9): 3538-3548.
- [8] FU Z Q, YANG Y, TU X T, et al. Learning a simple low-light image enhancer from paired low-light instances[C]. *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 17-24, 2023. *Vancouver, BC, Canada*. IEEE, 2023: 22252-22261.

- [9] WEI C, WANG W J, YANG W H, *et al.* Deep retinex decomposition for low-light enhancement t [EB/OL]. *arXiv*, 2018. <http://arxiv.org/abs/1808.04560>.
- [10] ZHANG Y H, ZHANG J W, GUO X J. Kindling the darkness: a practical low-light image enhancer [C]. *Proceedings of the 27th ACM International Conference on Multimedia. Nice France*. ACM, 2019: 1632-1640.
- [11] GUO C L, LI C Y, GUO J C, *et al.* Zero-reference deep curve estimation for low-light image enhancement [C]. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 13-19, 2020. Seattle, WA, USA. IEEE, 2020: 1777-1786.
- [12] JIANG Y F, GONG X Y, LIU D, *et al.* EnlightenGAN: deep light enhancement without paired supervision [J]. *IEEE Transactions on Image Processing*, 2021, 30: 2340-2349.
- [13] XU X G, WANG R X, FU C W, *et al.* SNR-aware low-light image enhancement [C]. *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 18-24, 2022. OrleansNew, LA, USA. IEEE, 2022: 17693-17703.
- [14] CAI Y H, BIAN H, LIN J, *et al.* Retinexformer: one-stage retinex-based transformer for low-light image enhancement [C]. *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*. Paris, France. IEEE, 2023: 12470-12479.
- [15] YU H, ZHENG N S, ZHOU M, *et al.* Frequency and Spatial dual guidance for Image dehazing [C]. *Computer Vision-ECCV 2022*. Cham: Springer Nature Switzerland, 2022: 181-198.
- [16] HUANG J, LIU Y J, ZHAO F, *et al.* Deep Fourier-based exposure correction network with Spatial-frequency interaction [C]. *Computer Vision-ECV 2022*. Cham: Springer Nature Switzerland, 2022: 163-180.
- [17] WANG C X, WU H J, JIN Z. FourLLIE: boosting low-light image enhancement by Fourier frequency information [C]. *Proceedings of the 31st ACM International Conference on Multimedia. Ottawa ON Canada*. ACM, 2023: 7459-7469.
- [18] AFIFI M, BROWN M S. Deep white-balance editing [C]. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 13-19, 2020. Seattle, WA, USA. IEEE, 2020: 1394-1403.
- [19] CUI Z, LI K, GU L, *et al.* You Only Need 90K Parameters to Adapt Light: a Light Weight Transformer for Image Enhancement and Exposure Correction [C]. *33rd British Machine Vision Conference 2022 (BMVC 2022)*. London, UK: BMVA Press, 2022.
- [20] GUO X J, HU Q M. Low-light image enhancement via breaking down the darkness [J]. *International Journal of Computer Vision*, 2023, 131 (1): 48-66.
- [21] ZAMIR S W, ARORA A, KHAN S, *et al.* Restormer: efficient transformer for high-resolution image restoration [C]. *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 18-24, 2022. OrleansNew, LA, USA. IEEE, 2022: 5718-5729.
- [22] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, *et al.* An image is worth 16×16 words: transformers for image recognition at scale [C]. *International Conference on Learning Representations (ICLR 2021)*. 2021.
- [23] VASWANI A, SHAZEER N, PARMAR N, *et al.* Attention is all you need [C]. *Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS 2017)*. Red Hook, NY, USA: Curran Associates Inc., 2017: 6000-6010.
- [24] WANG Z, BOVIK A C, SHEIKH H R, *et al.* Image quality assessment: from error visibility to structural similarity [J]. *IEEE Transactions on Image Processing*, 2004, 13(4): 600-612.
- [25] JOHNSON J, ALAHI A, LI F F. Perceptual losses for real-time style transfer and super-resolution [EB/OL]. *arXiv*, 2016. <http://arxiv.org/abs/1603.08155>.
- [26] SEIF G, ANDROUTSOS D. Edge-based loss function for single image super-resolution [C]. *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. April 15-20, 2018. Calgary, AB. IEEE, 2018: 1468-1472.
- [27] YANG W H, WANG W J, HUANG H F, *et al.* Sparse gradient regularized deep retinex network for robust low-light image enhancement [J]. *IEEE Transactions on Image Processing*, 2021, 30:

- 2072-2086.
- [28] LEE C, LEE C, KIM C S. Contrast enhancement based on layered difference representation of 2D histograms[J]. *IEEE Transactions on Image Processing*, 2013, 22(12): 5372-5384.
- [29] MA K D, ZENG K, WANG Z. Perceptual quality assessment for multi-exposure image fusion[J]. *IEEE Transactions on Image Processing*, 2015, 24(11): 3345-3356.
- [30] KINGMA D, BA J. Adam: A method for stochastic optimization[J]. *Computer Science*, 2014.
- [31] LOSHCIOLOV I, HUTTER F. SGDR: Stochastic Gradient Descent with Warm Restarts[C]. *Proceedings of the 5th International Conference on Learning Representations (ICLR 2017)*. Toulon, France: OpenReview.net, 2017.
- [32] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. *Communications of the ACM*, 2017, 60(6): 84-90.
- [33] MITTAL A, MOORTHY A K, BOVIK A C. No-reference image quality assessment in the spatial domain[J]. *IEEE Transactions on Image Processing*, 2012, 21(12): 4695-4708.
- [34] MITTAL A, SOUNDARARAJAN R, BOVIK A C. Making a 'completely blind' image quality analyzer[J]. *IEEE Signal Processing Letters*, 2013, 20(3): 209-212.
- [35] ZAMIR S W, ARORA A, KHAN S, et al. Learning enriched features for fast image restoration and enhancement[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023, 45(2): 1934-1948.
- [36] LIU R S, MA L, ZHANG J A, et al. Retinex-inspired unrolling with cooperative prior architecture search for low-light image enhancement[J]. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2021: 10556-10565.
- [37] WANG Y, WAN R, YANG W, et al. Low-Light Image Enhancement with Normalizing Flow[J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2022, 36(3): 2604-2612.
- [38] WANG T, ZHANG K H, SHEN T R, et al. Ultra-high-definition low-light image enhancement: a benchmark and transformer-based method[J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2023, 37(3): 2654-2662.
- [39] ZHOU S C, LI C Y, CHANG LOY C. LED-Net: joint low-light enhancement and Deblurring in the Dark [C]. *Computer Vision-ECCV 2022*. Cham: Springer Nature Switzerland, 2022: 573-589.
- [40] ZENG H, CAI J R, LI L D, et al. Learning image-adaptive 3D lookup tables for high performance photo enhancement in real-time[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020: 1.
- [41] YANG W, WANG S, FANG Y, et al. From fidelity to perceptual quality: a semi-supervised approach for low-light image enhancement[C]. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle, WA, USA: IEEE, 2020: 3060-3069.
- [42] ZHANG J, WANG X, LI Y, et al. PRNet: Low-Light Image Enhancement Based on Fourier Transform [J]. *IEEE Transactions on Instrumentation and Measurement*, 2025, 74: 1-14.

作者简介:



张易航(1999—),男,云南曲靖人,硕士研究生,主要从事图像处理、深度学习的研究。E-mail: 2023211521@stu.ppsuc.edu.cn

通讯作者:



钟 寒(1984—),女,山东济宁人,博士,副教授,硕士生导师,主要从事数据分析、计算机视觉和图像处理的研究。E-mail: zhonghan@ppsuc.edu.cn